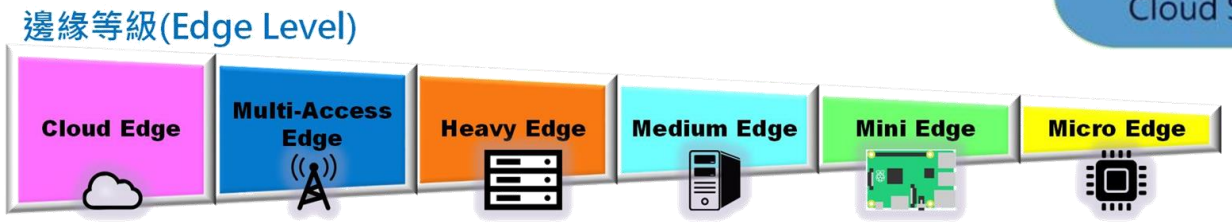
國立臺灣科技大學
NATIONAL TAIWAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Department of Computer Science & Information Engineering
資 訊 工 程 系

人工智慧與邊緣運算實務

7.4 邊緣智慧案例實作【影像分割】



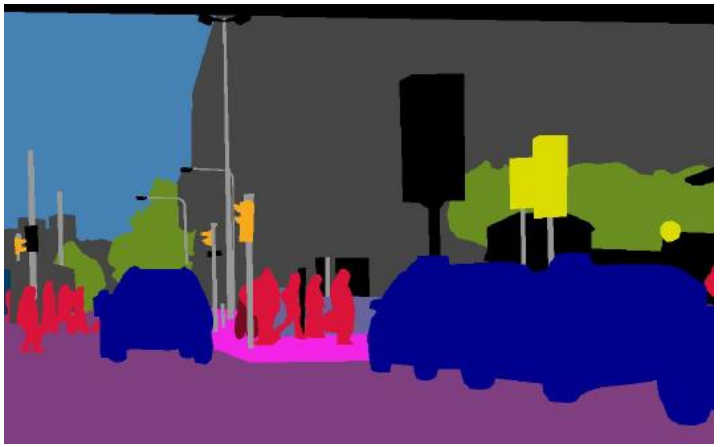
資訊工程系　許哲豪 助理教授

7.4 影像分割

- 影像分割技術
- 影像分割標註
- 語義分割範例
- 實例分割範例

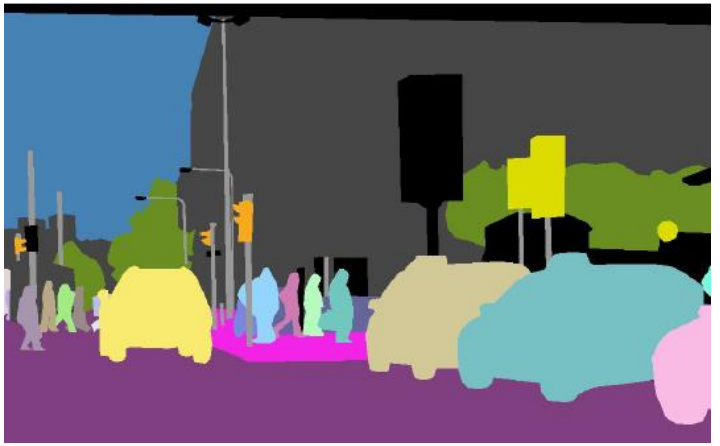
影像分割（像素級分類）類型

原始影像



語義 (Semantic)分割

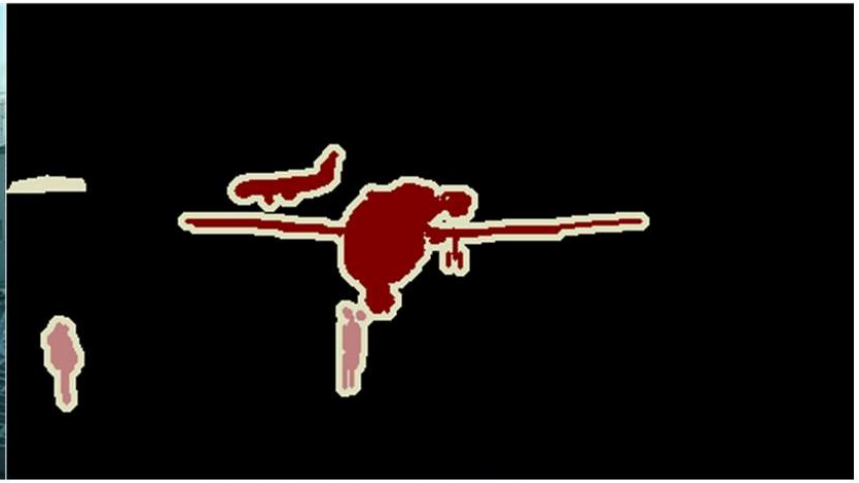
實例(Instance)分割



全景(Panoptic)分割

VOC語義分割資料格式

原始影像 (*.jpg)



語義分割後影像 (*.png)

黑色為背景

白色為邊界
（非必要）

PNG標註檔
採索引式
(Index)格式
類似GIF格式

Aeroplane	Diningtable	Bottle	Person	Chair	Train
Bicycle	Cat	Bus	Pottedplant	Cow	Tvmonitor
Bird	Horse	Car	Sheep		
Boat	Motorbike	Dog	Sofa		

VOC 20+1類別

COCO語義分割資料格式

Segmentation : [polygon] 封閉多邊形

```
"segmentation":          X1, Y1, X2,Y2, ... ,Xn, Yn
[[510.66,423.01,...,510.45,423.01]],
"area": 702.1057499999998,
"iscrowd": 0,
```

Segmentation : [RLE] 二進制影像壓縮

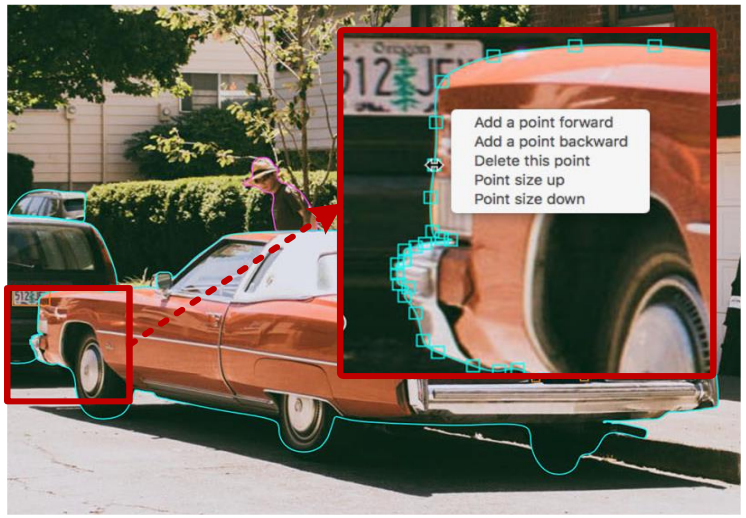
```
"segmentation": {"counts":
[20736,2,453,5,452,9,447,13,444,...,5,34552],
 "size": [457,640]
},
"area": 3074,
"iscrowd": 1,
```

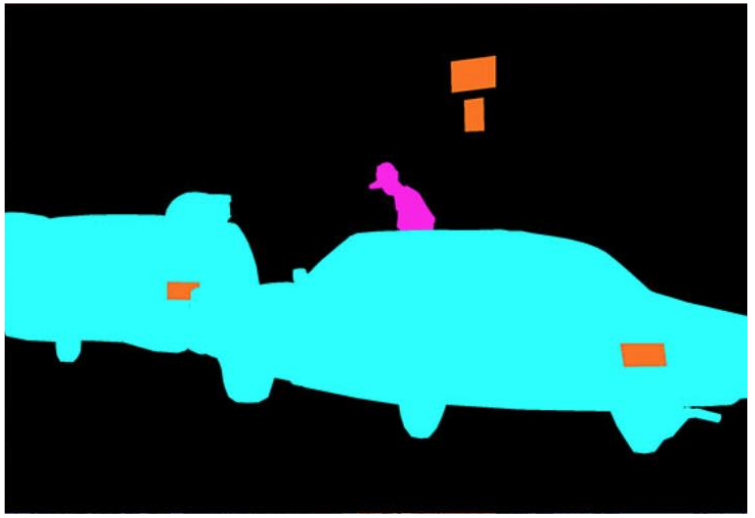
黑像素個數, 白像素個數, 黑, 白, ...

常見影像分割方式

貝茲曲線
(Bezier)

不受影像
尺寸影響





封閉多邊形
(Polygon)

不精準但
標註速度快

自由筆刷
(Brush)

很不精準

超像素
(Superpixel)

OpenCV
- LSC
- SEEDS
- SLIC

無自動合併

其它影像分割方式

- 特定封閉區域
  - 橢圓形、旋轉矩形、多角形...
- 特徵區域提取
  - 色彩、亮度、紋理...
  - 型態學分割、合併...
- 像素聚類
  - Seeds Growing, K-Mean, Mean-shift, Watershed, Grabcut…
- 其它客製軟體生成遮罩

影像分割標註工具

PixelAnnotationTool

https://github.com/abreheret/PixelAnnotationTool

基於OpenCV分水嶺方式實現語義分割，分割結果以圖檔輸出。

superpixels-segmentation

https://github.com/Labelbox/superpixels-segmentation

基於超像素SLIC方式實現影像分割，分割結果以圖檔輸出。

semantic-segmentation-editor

https://github.com/Hitachi-Automotive-And-Industry-Lab/semantic-segmentation-editor

基於WEB的標註工具，主要用於自駕車領域，支援2D(jpg, png)、3D(pcd)影像檔，具多邊形繪圖工具、魔術棒等工具，亦可對多邊形進行切割等工作。

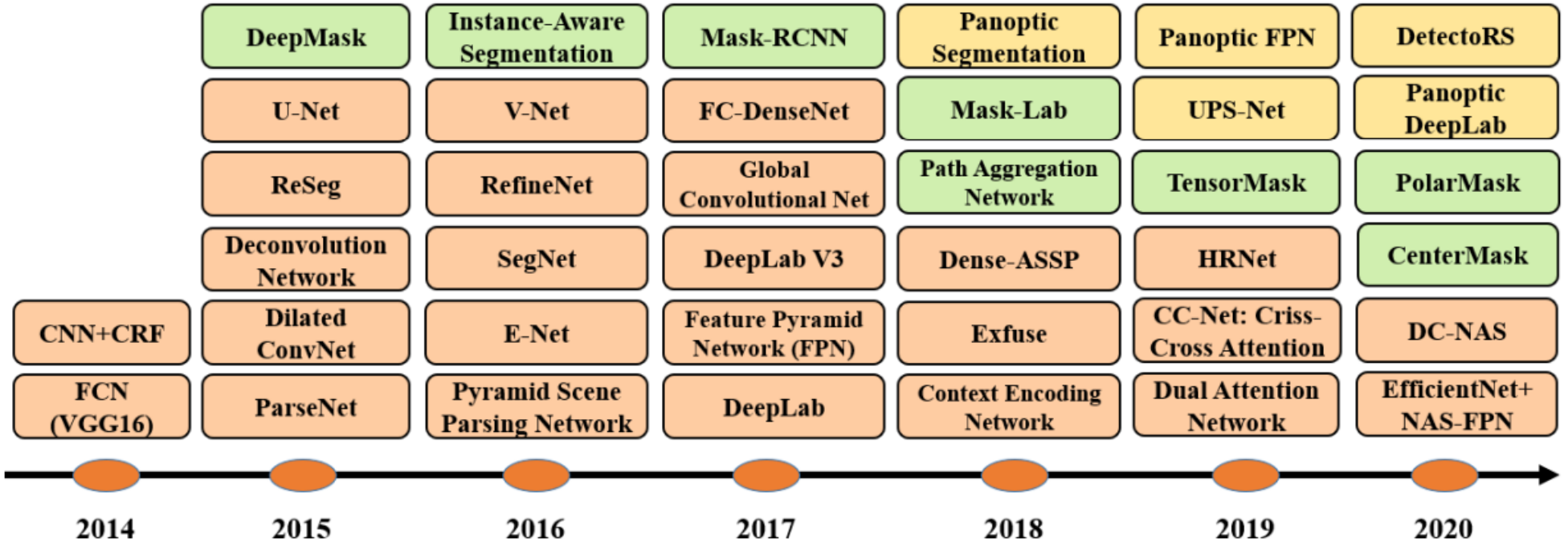
常見語義分割技術

1. Fully convolutional networks
2. Convolutional models with graphical models
3. Encoder-decoder based models
4. Multi-scale and pyramid network based models
5. R-CNN based models (for instance segmentation)
6. Dilated convolutional models and DeepLab family
7. Recurrent neural network based models
8. Attention-based models
9. Generative models and adversarial training
10. Convolutional models with active contour models
11. Other models

參考資料：https://arxiv.org/abs/2001.05566

深度學習影像分割技術演進

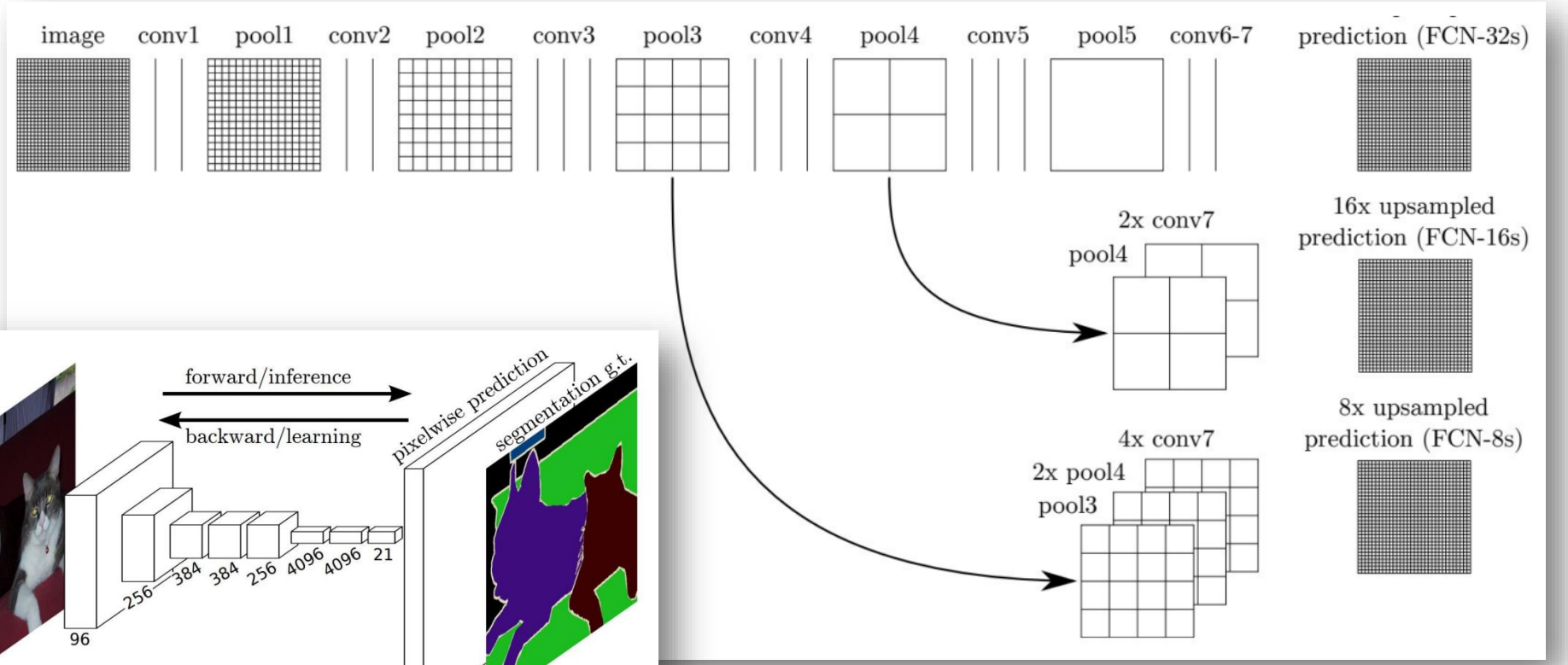
橙色：語義分割　　綠色：實例分割　　黃色：全景分割



參考資料：https://arxiv.org/abs/2001.05566

常見語義分割模型 FCN

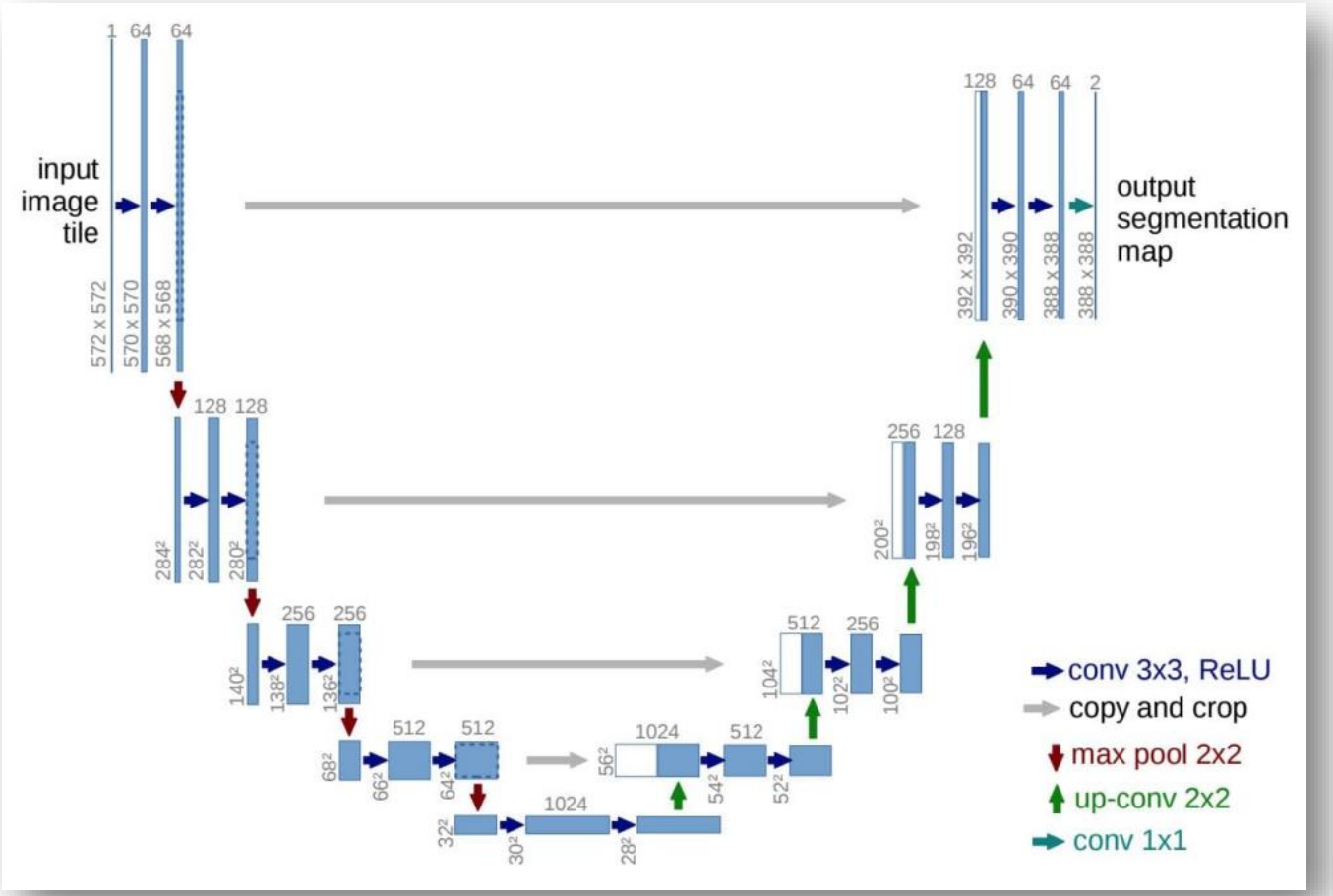
Fully Convolutional Networks Based Model

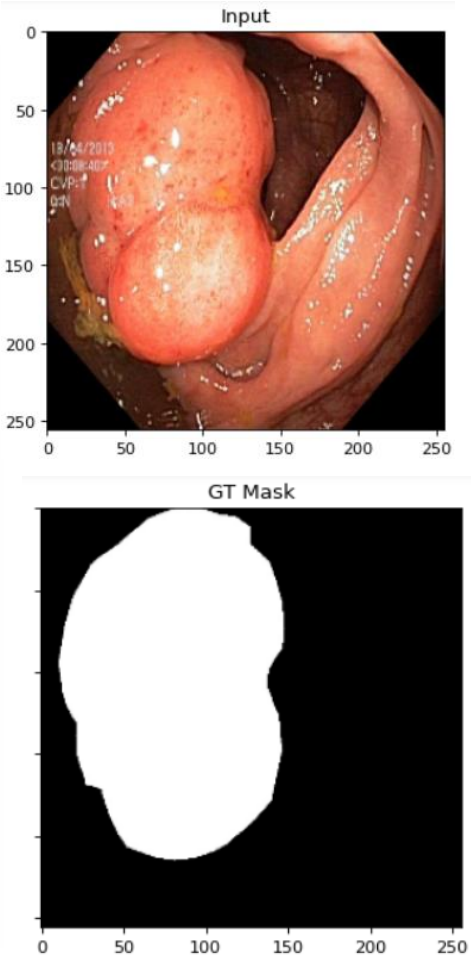


參考資料：https://arxiv.org/abs/2001.05566

常見語義分割模型 U-Net

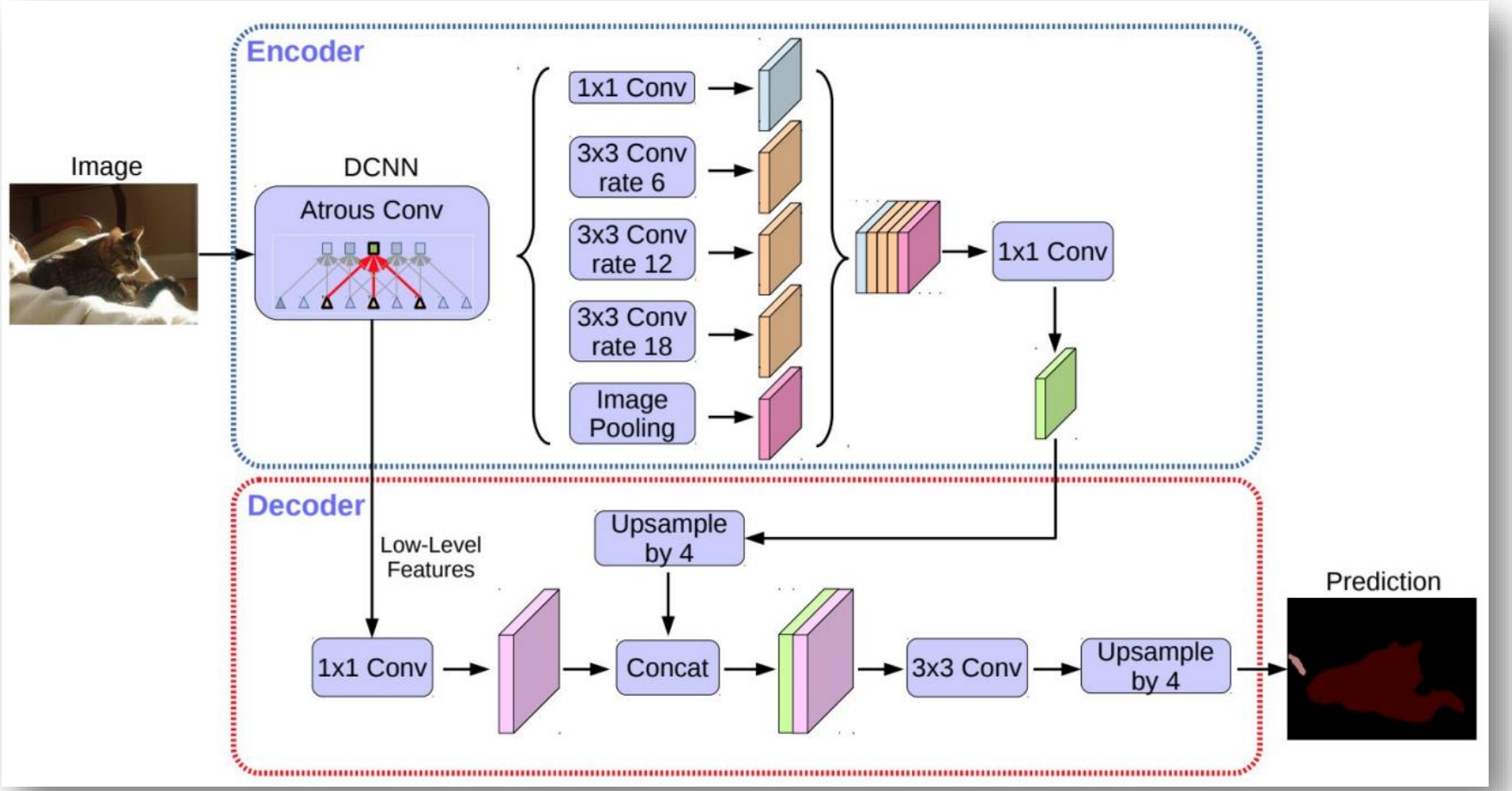
Encoder-Decoder Based Models





參考資料：https://arxiv.org/abs/2001.05566

常見語義分割模型 DeepLabV3+

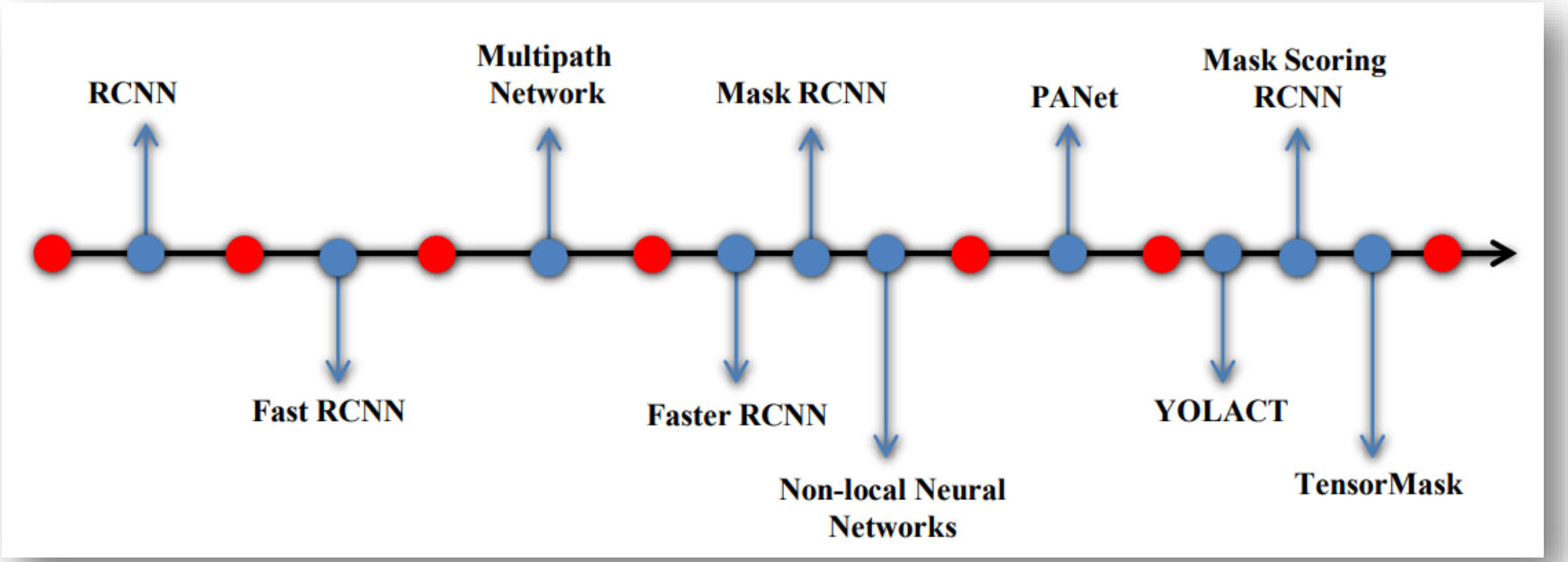


參考資料：https://arxiv.org/abs/2001.05566

常見實例分割技術

1. Classification of mask proposals
2. Detection followed by segmentation
3. Labelling pixels followed by clustering
4. Dense sliding window methods

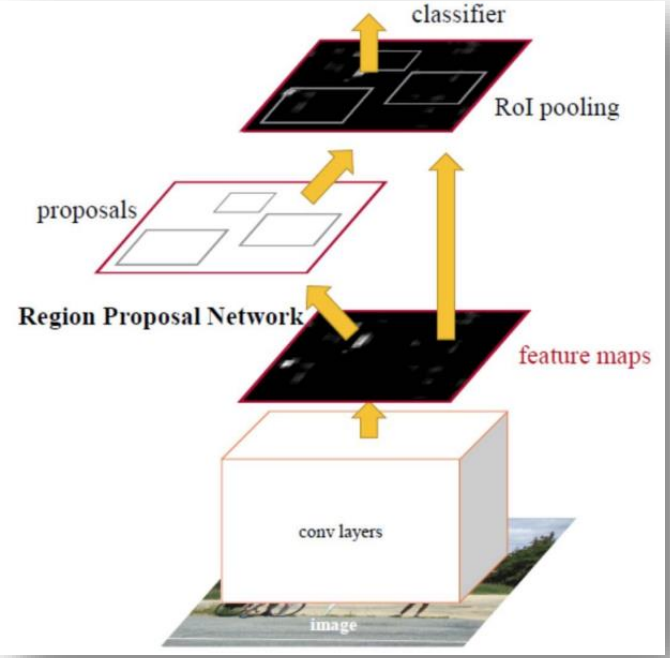
實例分割技術演進

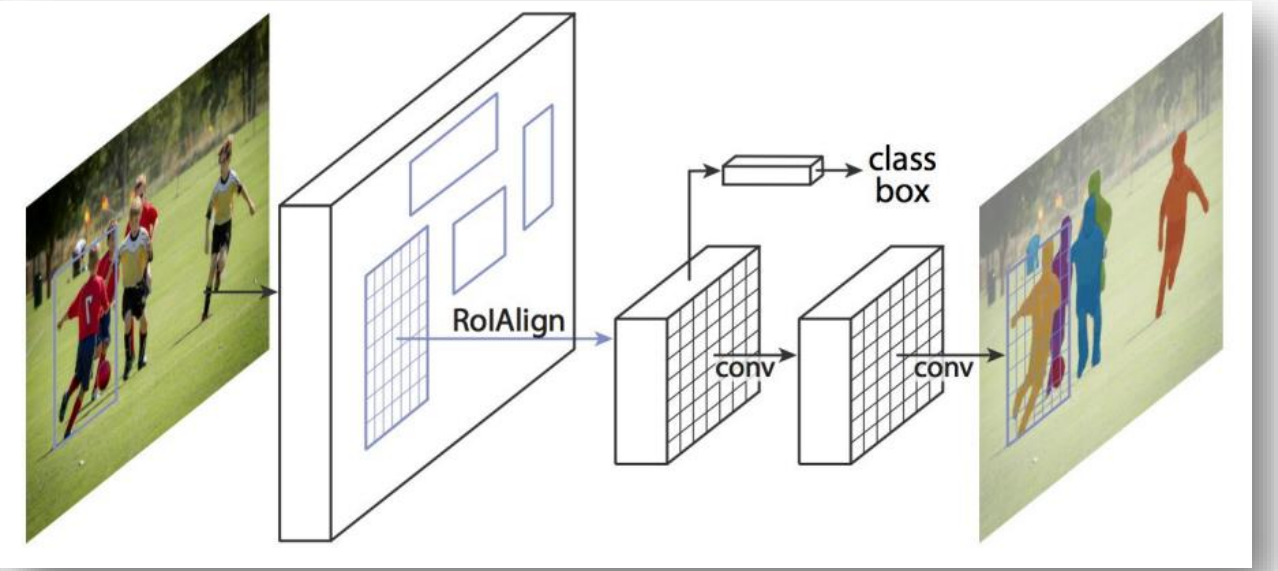


資料來源：https://arxiv.org/abs/2007.00047

常見實例分割模型 Mask R-CNN

Region Convolution Neural Network Based Model



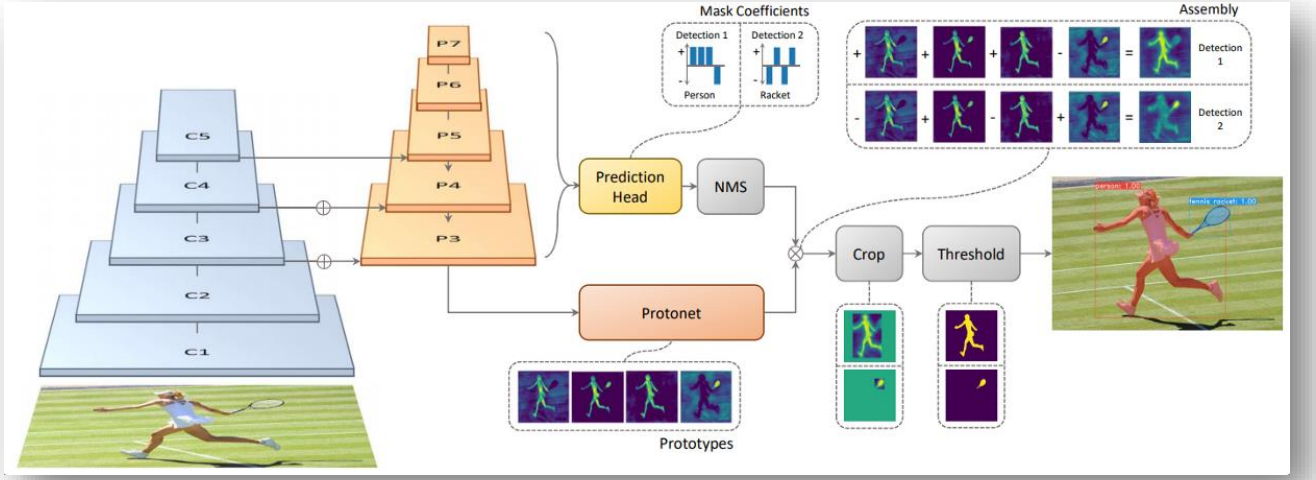


Faster R-CNN

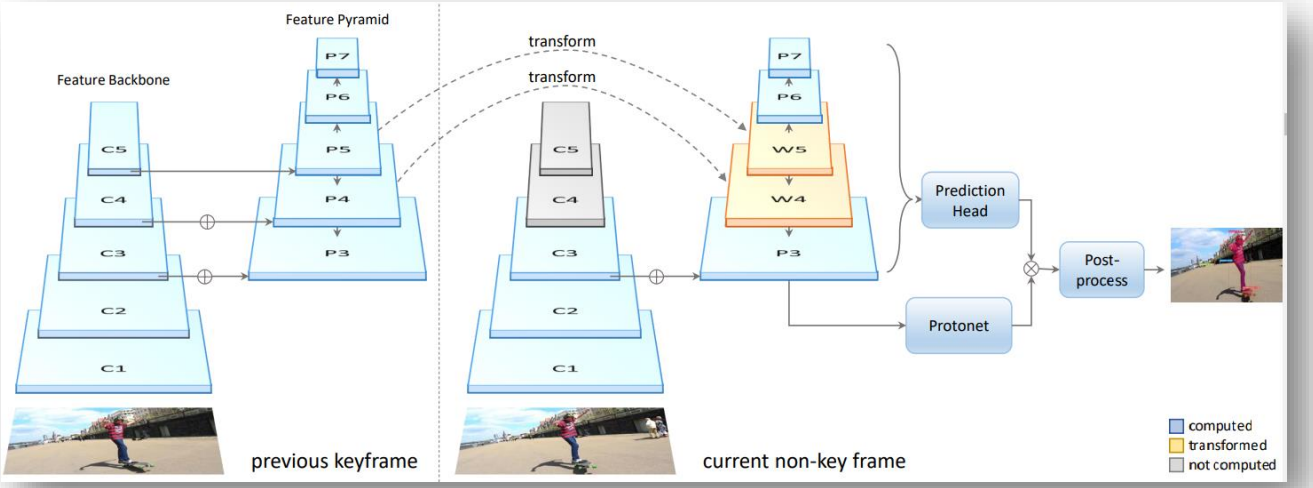
Mask R-CNN

參考資料：https://arxiv.org/abs/2001.05566

常見實例分割模型 Yolact & Edge



You Only Look At CoefficienTs

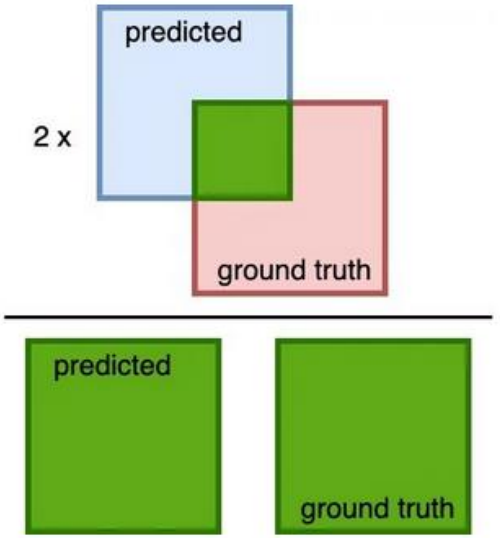


YOLACT Edge

評量指標 DICE Coefficient

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

Dice coefficient = $\frac{\text{2x area of overlapped (green)}}{\text{total area (green)}}$ = $\frac{\text{2 x predicted} \cap \text{ground truth}}{\text{predicted} + \text{ground truth}}$



Intersection over Union

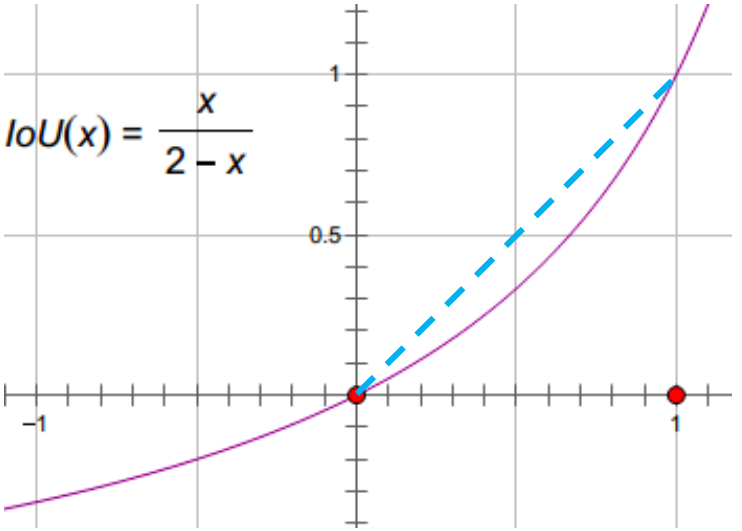
預測與真實
交集 / 聯集

$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)}$$

predicted　　Ground truth

真陽：True Positive 真陰：True Negative
偽陽：False Positive 偽陰：False Negative

$$IoU(x) = \frac{x}{2 - x}$$



OpenVINO影像分割預訓練模型

- **Intel's Pre-trained**
  - Semantic Segmentation Models 語義分割
  - Instance Segmentation Models 實例分割
- **Public Pre-trained**
  - Semantic Segmentation Models
  - Instance Segmentation Models
  - 3D Semantic Segmentation Models

- **Demos**
  - Image Segmentation Python*
  - Instance Segmentation Python*

Intel's Pretrained語義分割模型

Model Name	Size	GFlops	MPara.	Mean IOU	Classes
road-segmentation-adas-0001	896x512	4.770	0.184	84.40%	4
semantic-segmentation-adas-0001	2048x1024	58.572	6.686	69.07%	20
unet-camvid-onnx-0001	368x480	260.1	31.03	71.95%	12
icnet-camvid-ava-0001	720x960	75.8180	26.7043	75.42%	12
icnet-camvid-ava-sparse-30-0001	720x960	75.8180	26.7043	75.87%	12
icnet-camvid-ava-sparse-60-0001	720x960	75.8180	26.7043	75.79%	12

適用範例程式：
Image Segmentation C++ DemoImage、Segmentation Python* Demo

資料來源： https://docs.openvino.ai/latest/omz_models_group_intel.html#semantic-segmentation-models

Public Pretrained語義分割模型

Model Name	Size	GFlops	MPara.	Mean IOU	Classes
deeplabv3	513x513	11.469	23.819	68.41%	21 (VOC)
drn-d-38	1024x2048	1768.32	25.9939	71.31%	30 (Cityscapes)
hrnet-v2-c1-segmentation	320x320	81.993	66.4768	33.02%	(ADE20K)
fastseg-large	1024x2048	140.961	3.2	72.67%	19 (Cityscapes)
fastseg-small	1024x2048	69.220	1.1	67.15%	19 (Cityscapes)
pspnet-pytorch	512x512	357.171	46.5827	70.1%	21 (VOC)
ocrnet-hrnet-w48-paddle	1024x2048	324.66	70.47	82.15%	19 (Cityscapes)

適用範例程式：
Image Segmentation C++ DemoImage、Segmentation Python* Demo

資料來源： https://docs.openvino.ai/latest/omz_models_group_public.html#segmentation-models

Demos語義分割範例

- Image Segmentation Python* Demo

Liunx (含Ubuntu, DevCloud)下範例路徑
/opt/intel/openvino_2021/deployment_tools/open_model_zoo/demos/segmentation_demo/python/segmentation_demo.py

延遲(Latency)包括
- 解碼(Decode)
- 預處理(Preprocessing)
- 推論(Inference)
- 後處理(Postprocessing)
- 渲染(Rendering)

每秒影格數(FPS)
Frame per Second

資料來源：https://docs.openvino.ai/latest/omz_demos_segmentation_demo_python.html

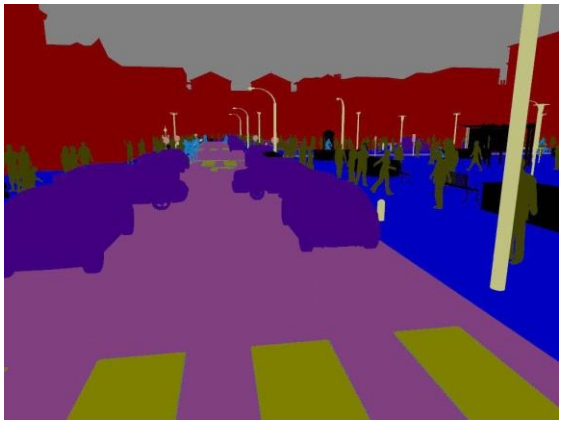
語義分割範例可支援模型

- architecture_type = segmentation
  - Deeplabv3
  - drn-d-38
  - fastseg-large
  - fastseg-small
  - hrnet-v2-c1-segmentation
  - icnet-camvid-ava-0001
  - icnet-camvid-ava-sparse-30-0001
  - icnet-camvid-ava-sparse-60-0001
  - ocrnet-hrnet-w48-paddle
  - pspnet-pytorch
  - road-segmentation-adas-0001
  - semantic-segmentation-adas-0001
  - unet-camvid-onnx-0001
- architecture_type = salient_object_detection
  - f3net

資料來源：https://docs.openvino.ai/latest/omz_demos_segmentation_demo_python.html

範例7-4-1 語義分割(Intel's)

輸入影像



標註影像

輸出影像

模型名稱：

semantic-segmentation-adas-0001

分類名稱：

road, sidewalk, building, wall, fence, pole, traffic light, traffic sign, vegetation, terrain, sky, person, rider, car, truck, bus, train, motorcycle, bicycle, ego-vehicle

主要用途：

使用Intel預訓練語義分割模型進行影像分割並著色。

資料來源：https://docs.openvino.ai/latest/omz_models_model_semantic_segmentation_adas_0001.html#doxid-omz-models-model-semantic-segmentation-adas-0001

範例7-4-2 語義分割(Public)

輸入影像

輸出影像

模型名稱：

deeplabv3

分類名稱：(VOC dataset)

background, person, bird, cat, cow, dog, horse, sheep, aeroplane, bicycle, boat, bus, car, motorbike, train, bottle, chair, dining table, potted plant, sofa, tv/monitor

主要用途：

使用Public預訓練語義分割模型，以OpenVINO轉換優化後進行影像分割並著色。

資料來源：https://docs.openvino.ai/latest/omz_models_model_deeplabv3.html#doxid-omz-models-model-deeplabv3

Intel's Pretrained實例分割模型

- instance-segmentation-security-xxxx（左圖）
  - xxxx = 0002 / 0091 / 0228 / 1039 / 1040
  - COCO Dataset 80 分類

- instance-segmentation-person-xxxx（右圖）
  - xxxx = 0007
  - COCO Dataset 80 分類

（限OpenVINO 2022.1後使用）

資料來源：https://docs.openvino.ai/latest/omz_models_group_intel.html#instance-segmentation-models

Intel's Pretrained實例分割模型

Model Name	Size	GFlops	MPara.	Mask AP	Classes
instance-segmentation-security-0002	1024x768	423.0842	48.3732	36.44%	21 (VOC)
instance-segmentation-security-0091	1344x800	828.6324	101.236	38.14%	21 (VOC)
instance-segmentation-security-0228	608x608	147.2352	49.8328	33.9%	21 (VOC)
instance-segmentation-security-1039	480x480	13.9672	10.5674	28.6%	21 (VOC)
instance-segmentation-security-1040	608x608	29.334	13.5673	31.2%	21 (VOC)
instance-segmentation-person-0007	320x544	4.8492	7.2996	30.9%	1 (VOC)

適用範例程式：

Instance Segmentation Python* Demo Background Subtraction Python （限OpenVINO 2022.1後使用）

資料來源： https://docs.openvino.ai/latest/omz_models_group_intel.html#semantic-segmentation-models

Public Pretrained實例分割模型

Model Name	Size	GFlops	MPara.	Mask AP	Classes
mask_rcnn_inception_resnet_v2_atrous_coco	1365x800	675.314	92.368	35.36%	21 (VOC)
mask_rcnn_resnet50_atrous_coco	1365x800	294.738	50.222	27.46%	21 (VOC)
yolact-resnet50-fpn-pytorch	550x550	118.575	36.829	30.69%	21 (VOC)

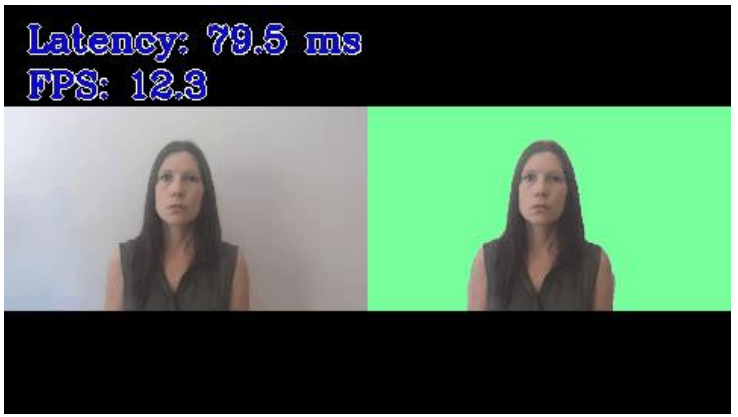
註：前兩項因模型太大，只支援CPU及iGPU，不支援MYRIAD（NCS2）

適用範例程式：
Instance Segmentation Python* Demo

資料來源：https://docs.openvino.ai/latest/omz_models_group_public.html#instance-segmentation-models

Demos實例分割範例

- Liunx (含Ubuntu, DevCloud)下範例路徑 /opt/intel/openvino_2021/deployment_tools/open_model_zoo/demos/instance_segmentation_demo/python/
  - instance_segmentation_demo.py
  - background_subtraction_demo.py（限OpenVINO 2022.1後使用）



資料來源：https://docs.openvino.ai/latest/omz_demos.html

範例7-4-3實例分割(Intel's)

輸入影像

輸出影像

模型名稱：

instance-segmentation-security-0228

分類名稱：

COCO Dataset 80 Classes

主要用途：

使用Intel's預訓練實例分割模型，以OpenVINO轉換優化後進行影像分割並著色，相同物件（實例）會著不同顏色，可支援標籤顯示。

資料來源：https://docs.openvino.ai/latest/omz_models_model_instance_segmentation_security_0228.html

範例7-4-4 影像去背 U² (Notebook)

輸入影像

輸出影像

輸出影像

輸出影像

模型名稱：
U^2-Net

分類名稱：
無

主要用途：
使用預訓練影像分割模型，配合OpenVINO進行影像分割找出前景及背景，可選擇性替換其它背景。

OpenVINO Tutorials Notebooks
205-vision-background-removal

參考資料：https://docs.openvino.ai/latest/notebooks/205-vision-background-removal-with-output.html

參考文獻

- 許哲豪，【AI HUB專欄】如何建立精準標註的電腦視覺資料集

https://omnixri.blogspot.com/2020/10/ai-hub_16.html

- 許哲豪，【OpenVINO™教學】土炮影像實例分割型智慧監控系統

https://omnixri.blogspot.com/2019/09/openvino.html

- Shervin Minaee etc., Image Segmentation Using Deep Learning: A Survey

https://arxiv.org/abs/2001.05566

- Abdul Mueed Hafiz etc., A Survey on Instance Segmentation: State of the art

https://arxiv.org/abs/2007.00047